Linear Algebra and its Applications 436 (2012) 3777–3783

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Unitary similarity to a normal matrix

Tatiana G. Gerasimova

Faculty of Mechanics and Mathematics, Kiev National Taras Shevchenko University, Kiev, Ukraine

ARTICLE INFO

Article history:
Received 17 November 2011
Accepted 20 November 2011
Available online 11 January 2012

Submitted by R. Brualdi

AMS classification:
15A21
15A60
47B15

Keywords:
Unitary similarity
Normal matrix
Frobenius norm
Spectral norm

ABSTRACT

We give several criteria of unitary similarity of a normal matrix A and any matrix B in terms of the Frobenius and spectral norms, characteristic polynomials, and traces of matrices.

1. Introduction

A classical problem in linear algebra is the following one: how can one determine whether square complex matrices A and B are unitarily similar (i.e., $U^{-1}AU = B$ for a unitary U)? The most known solution is Specht's theorem [10]: A and B are unitarily similar if and only if

$$\operatorname{tr}\ \omega(A, A^*) = \operatorname{tr}\ \omega(B, B^*)$$

for all words ω in two noncommuting variables, in which $\operatorname{tr} A$ denotes the trace of A.

Suppose that A is an upper triangular Toeplitz matrix with nonzero superdiagonal and B is any of the same size. Then A and B are unitarily similar if and only if the matrices $f(A)$ and $f(B)$ have the same Frobenius or spectral norm for all complex polynomials f (see [3, Theorem 2.1,4, Theorem 2.1]). We prove analogous statements for a normal matrix A and any B. Recall that the Frobenius and spectral

E-mail address: gerasimova@imath.kiev.ua

0024-3795/$ - see front matter
doi:10.1016/j.laa.2011.11.040

norms of a complex matrix $C = [c_{ij}]$ are defined as follows:

$$\|C\| := \sqrt{\sum |c_{ij}|^2}, \quad \|C\|_{\mathrm{sp}} := \max_{|v|=1} |Cv|$$

in which $|\cdot|$ is the Euclidean norm of vectors.

In the process of proving we establish that a normal A and any B are unitarily similar if and only if $\|A\| = \|B\|$ and $\operatorname{tr} A^k = \operatorname{tr} B^k$ for $k = 1, \ldots, n$. A similar statement was proved by Murnaghan [9] and, independently, by Ikramov [8]: two normal matrices A and B are unitarily similar if and only if $\operatorname{tr} A^k = \operatorname{tr} B^k$ for $k = 1, \ldots, n$ (note that this statement does not follow directly from Specht's theorem, see [8]).

Our goal is to prove the following theorem in which we give several criteria of unitary similarity of a normal A and any B. Matrices that are unitarily similar to a symmetric matrix are studied in [1,5].

Theorem. *Let A be an $n \times n$ normal complex matrix and B be any $n \times n$ complex matrix. The following statements are equivalent:*

- (i) *A and B are unitarily similar;*
- (ii) *B is normal (i.e., B satisfies one of 89 criteria of normality from [2,6]) and the characteristic polynomials of A and B are equal;*
- (iii) *$\|A\| = \|B\|$ and the characteristic polynomials of A and B are equal;*
- (iv) *$\|A\| = \|B\|$ and $\operatorname{tr} A^k = \operatorname{tr} B^k$ for $k = 1, \ldots, n$;*
- (v) *$\|A^k + cI_n\| = \|B^k + cI_n\|$ for $c \in \{0, 1, i\}$ and $k = 1, \ldots, n$;*
- (vi) *$\|f(A)\| = \|f(B)\|$ for all $f \in \mathbb{C}[x]$ of degree at most n;*
- (vii) *$\|f(A)\|_{\mathrm{sp}} = \|f(B)\|_{\mathrm{sp}}$ for all $f \in \mathbb{C}[x]$ of degree at most n, and the characteristic polynomials of A and B are equal.*

The condition (i) implies (ii)–(vii) since the Frobenius and spectral norms of a matrix do not change under multiplication by unitary matrices. The implication (ii)⇒(i) is obvious. The implications (iv)⇒(iii)⇒(i), (vi)⇒(v)⇒(iv), and (vii)⇒(i) are proved in Sections 2, 3, and 4, respectively.

Remark.

(a) If B is normal, then the condition $\|A\| = \|B\|$ in (iv) can be omitted (see [8] or [9]). It cannot be omitted for an arbitrary B: if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

then they have the same characteristic polynomial and $\operatorname{tr} A^i = \operatorname{tr} B^i$ for $i = 1, \ldots, n$, but A and B are not unitarily similar.

(b) If B is normal, then the condition "$\|A^k + cI_n\| = \|B^k + cI_n\|$ for $c \in \{0, 1, i\}$" in (v) can be replaced by $\|A^k\| = \|B^k\|$. It cannot be replaced for an arbitrary B: if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

then $\|A^k\| = \|B^k\|$ for $k = 1, \ldots, n$, but A and B are not unitarily similar.

(c) The condition $\|f(A)\|_{\mathrm{sp}} = \|f(B)\|_{\mathrm{sp}}$ in (vii) cannot weaken to $\|A\|_{\mathrm{sp}} = \|B\|_{\mathrm{sp}}$. For example, if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

then they have the same characteristic polynomial and $\|A\|_{\mathrm{sp}} = \|B\|_{\mathrm{sp}} = 2$, but A and B are not unitarily similar.

(d) The condition "the characteristic polynomials of A and B are equal" in (vii) cannot be omitted. For example, if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

then $\|f(A)\|_{\mathrm{sp}} = \|f(B)\|_{\mathrm{sp}} = \max(|f(1)|, |f(2)|)$ for all $f \in \mathbb{C}[x]$, but A and B are not unitarily similar.

2. Proof of the implications (iv) ⇒ (iii) ⇒ (i)

Let A be an $n \times n$ normal matrix and let B be any matrix of the same size.

First we prove (iii) ⇒ (i). Suppose that A and B satisfy (iii). Since A and B have the same characteristic polynomials, they have the same eigenvalues. Using transformations of unitary similarity (which preserve (iii)), we first reduce A to diagonal form and then B to upper triangular form with the same main diagonal,

$$A = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}, \quad B = \begin{bmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & b_{n-1,n} \\ 0 & & & \lambda_n \end{bmatrix},$$

which is possible by [7, Theorem 2.3.1].

Since $\|A\| = \|B\|$,

$$\sum_{i<j} |b_{ij}|^2 = 0.$$

Thus, B is diagonal and is equal to A up to permutation of diagonal entries. Hence A and B are unitary similar, which proves (iii) ⇒ (i).

By [7, Problem 12 in Section 1.2], Newton's identities ensure that two $n \times n$ matrices A and B have the same characteristic polynomial if and only if $\operatorname{tr} A^k = \operatorname{tr} B^k$ for $k = 1, \dots, n$, which proves (iv) ⇒ (iii).

3. Proof of the implication (vi) ⇒ (v) ⇒ (iv)

Let A be an $n \times n$ normal matrix and let B be a matrix of the same size. The implication (vi) ⇒ (v) is obvious, let us prove (v) ⇒ (iv).

Suppose that A and B satisfy (v). Without loss of generality, we assume that

$$A=\begin{bmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{bmatrix},\quad B=\begin{bmatrix}\mu_1 & b_{12} & \dots & b_{1n}\\ & \mu_2 & \ddots & \vdots\\ & & \ddots & b_{n-1,n}\\ 0 & & & \mu_n\end{bmatrix}$$

(due to [7, Theorems 2.5.3, 2.3.1] we can reduce them to this form by transformations of unitary similarity, which preserve (v)).

Denote by $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ the real and imaginary parts of $z \in \mathbb{C}$. The condition (v) implies $\|A\|^2 = \|B\|^2$ and $\|A+I\|^2 = \|B+I\|^2$, i.e.

$$\sum |\lambda_i|^2 = \sum |\mu_i|^2 + \sum_{i<j} |b_{ij}|^2 \tag{1}$$

and

$$\sum |\lambda_i + 1|^2 = \sum |\mu_i + 1|^2 + \sum_{i<j} |b_{ij}|^2. \tag{2}$$

Subtracting (1) from (2), we obtain

$$\sum(|\lambda_i + 1|^2 - |\lambda_i|^2) = \sum(|\mu_i + 1|^2 - |\mu_i|^2),$$

$$\sum((\lambda_i + 1)(\overline{\lambda_i} + 1) - \lambda_i\overline{\lambda_i}) = \sum((\mu_i + 1)(\overline{\mu_i} + 1) - \mu_i\overline{\mu_i}),$$

$$\sum(\lambda_i + \overline{\lambda_i} + 1) = \sum(\mu_i + \overline{\mu_i} + 1),$$

$$\sum \mathrm{Re}\,\lambda_i = \sum \mathrm{Re}\,\mu_i. \tag{3}$$

By (v), $\|A + iI\|^2 = \|B + iI\|^2$, and so

$$\sum |\lambda_i + i|^2 = \sum |\mu_i + i|^2 + \sum_{i<j} |b_{ij}|^2. \tag{4}$$

Subtracting (1) from (4), we obtain

$$\sum \mathrm{Im}\,\lambda_i = \sum \mathrm{Im}\,\mu_i. \tag{5}$$

By (3) and (5),

$$\sum \lambda_i = \sum \mu_i, \tag{6}$$

and so $\mathrm{tr}\,A = \mathrm{tr}\,B$. The same reasoning to A^k and B^k instead of A and B ensure that

$$\mathrm{tr}\,A^k = \mathrm{tr}\,B^k \quad \text{for all } k = 1, \dots, n, \tag{7}$$

which proves (v)$\Rightarrow$(iv).

4. Proof of the implication (vii) ⇒ (i)

Let A be an $n \times n$ normal matrix and let B be any matrix of the same size. Suppose that A and B satisfy (vii).

Since A and B have the same characteristic polynomial, they have the same eigenvalues. Using transformations of unitary similarity (which preserve (vii)), we reduce A to diagonal form, B to upper triangular form with the same main diagonal, and obtain

$$A = \begin{bmatrix} \lambda_1 I_{m_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_s I_{m_s} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & & * \\ & \ddots & \\ 0 & & B_s \end{bmatrix},$$

in which $\lambda_i \neq \lambda_j$ if $i \neq j$ and every B_i is an $m_i \times m_i$ upper triangular matrix with the main diagonal $(\lambda_i, \ldots, \lambda_i)$.

Let us prove that

$$B_1 = \lambda_1 I_{m_1}, \ \ldots, \ B_s = \lambda_s I_{m_s}. \tag{8}$$

The minimal polynomial of A is $\mu_A(x) = (x - \lambda_1) \cdots (x - \lambda_s)$. Since $\mu_A(A) = 0$, by (vii) we have

$$\mu_A(B) = \begin{bmatrix} \mu_A(B_1) & & * \\ & \ddots & \\ 0 & & \mu_A(B_s) \end{bmatrix} = 0,$$

and so for each i

$$\mu_A(B_i) = (B_i - \lambda_1 I) \cdots (B_i - \lambda_s I) = 0.$$

But $\det(B_i - \lambda_j I) \neq 0$ if $i \neq j$, hence all $B_i - \lambda_i I = 0$, which ensures (8).

We need to prove that $B = B_1 \oplus \cdots \oplus B_s$. Assume to the contrary that B is not diagonal. Let l be the maximum index such that there is a nonzero entry over $B_l = \lambda_l I_{m_l}$ in B.

If $l < s$, then B has the form

$$B = \begin{bmatrix} C & Y \\ 0 & B_l \end{bmatrix} \oplus D,$$

in which C is upper triangular, $Y \neq 0$, and $D = B_{l+1} \oplus \cdots \oplus B_s$. By permutations of rows and the same permutations of columns simultaneously in A and B, we make

$$B = D \oplus \begin{bmatrix} C & Y \\ 0 & B_l \end{bmatrix}.$$

Thus, we can suppose that $l = s$, then

$$B = \begin{bmatrix} C & Y \\ 0 & \lambda_s I_{m_s} \end{bmatrix},$$

in which $Y \neq 0$ and

$$C = \begin{bmatrix} \lambda_1 I_{m_1} & & * \\ & \ddots & \\ 0 & & \lambda_{s-1} I_{m_{s-1}} \end{bmatrix}. \tag{9}$$

The minimal polynomial of C is $\mu_C(x) := (x - \lambda_1) \cdots (x - \lambda_{s-1})$, and so

$$\mu_C(A) = \begin{bmatrix} 0_{m'} & 0 \\ 0 & \mu_C(\lambda_s) I_{m_s} \end{bmatrix}, \quad \mu_C(B) = \begin{bmatrix} 0_{m'} & Z \\ 0 & \mu_C(\lambda_s) I_{m_s} \end{bmatrix},$$

in which $m' := m_1 + \cdots + m_{s-1}$ and

$$Z = \sum_{i=1}^{s-1} (C - \lambda_1 I) \cdots (C - \lambda_{i-1} I) Y (\lambda_s - \lambda_{i+1}) \cdots (\lambda_s - \lambda_{s-1}).$$

Thus, $Z = f(C)Y$ with

$$f(x) := \sum_{i=1}^{s-1} (x - \lambda_1) \cdots (x - \lambda_{i-1})(\lambda_s - \lambda_{i+1}) \cdots (\lambda_s - \lambda_{s-1}).$$

By (vii),

$$\left\| \begin{bmatrix} 0_{m'} & Z \\ 0 & \mu_C(\lambda_s) I_{m_s} \end{bmatrix} \right\| = \|\mu_C(B)\| = \|\mu_C(A)\| = |\mu_C(\lambda_s)|,$$

hence $Z = 0$ and $f(C)Y = 0$. Since $Y \neq 0$, $f(C)$ is a singular matrix and so $f(\lambda_1) f(\lambda_2) \cdots f(\lambda_{s-1}) = 0$ by (9). Therefore,

$$f(\lambda_r) = 0 \quad \text{for some } r \leqslant s - 1. \tag{10}$$

Write

$$M := \begin{bmatrix} \lambda_r & 1 \\ 0 & \lambda_s \end{bmatrix}.$$

Equating the $(1, 2)$ entries in the matrices

$$(M - \lambda_1 I)(M - \lambda_2 I) \cdots (M - \lambda_{s-1} I) = (M - \lambda_r I)(M - \lambda_1 I) \cdots (M - \lambda_{r-1} I)(M - \lambda_{r+1} I) \cdots (M - \lambda_{s-1} I),$$

we obtain

$$f(\lambda_r) = (\lambda_s - \lambda_1) \cdots (\lambda_s - \lambda_{r-1})(\lambda_s - \lambda_{r+1}) \cdots (\lambda_s - \lambda_{s-1}) \neq 0,$$

which contradicts (10).

References

[1] L. Balayan, S.R. Garcia, Unitary equivalence to a complex symmetric matrix: geometric criteria, Oper. Matrices 4 (2010) 53–76.
[2] L. Elsner, Kh.D. Ikramov, Normal matrices: an update, Linear Algebra Appl. 285 (1998) 291–303.
[3] D. Farenick, V. Futorny, T.G. Gerasimova, V.V. Sergeichuk, N. Shvai, Criterion of unitary similarity for upper triangular matrices in general position, Linear Algebra Appl. 435 (2011) 1356–1369.
[4] D. Farenick, T.G. Gerasimova, N. Shvai, A complete unitary similarity invariant for unicellular matrices, Linear Algebra Appl. 435 (2011) 409–419.
[5] S.R. Garcia, D.E. Poore, M.K. Wyse, Unitary equivalence to a complex symmetric matrix: a modulus criterion, Oper. Matrices 5 (2011) 273–287.
[6] R. Grone, C.R. Johnson, E.M. Sa, H. Wolkowicz, Normal matrices, Linear Algebra Appl. 87 (1987) 213–225.
[7] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[8] Kh.D. Ikramov, On the unitary similarity of normal matrices from the standpoint of Specht's criterion, Comput. Math. Math. Phys. 42 (2002) 1065–1066.
[9] F.D. Murnaghan, On the unitary invariants of a square matrix, An. Acad. Brasil. Ciênc. 26 (1954) 1–7.
[10] W. Specht, Zur Theorie der Matrizen. II, Jahresber. Deutsch. Math.-Verein. 50 (1940) 19–23.